

A Scoping Review of the Use of Machine Learning in Health Economics and Outcomes Research : Data from Wearable Devices

Woojung Lee, PharmD, Naomi Schwartz, PhD, Aasthaa Bansal, PhD, Sara Khor, MASc, Noah Hammarlund, PhD, Anirban Basu, PhD, Beth Devine, PharmD, PhD
The Comparative Health Outcomes, Policy, and Economics (CHOICE) Institute, School of Pharmacy, University of Washington

THE CHOICE INSTITUTE

School of Pharmacy



Background

- > With the recent advancement in wearable technology, the market for wearable devices has been rapidly growing
- > A continuous stream of individual data from wearable devices
- > Machine learning (ML) has been gaining increasing attention in managing the generation of large amounts of wearable data
- > The adoption of wearable data in HEOR has been slow and there are uncertainties in how and when to use ML with wearable data



Objective

- > **To review the emerging patterns of machine learning applications on wearable data in health economics and outcome research**
 - Can help researchers understand potential ways of using ML with wearable data



Methods



THE CHOICE INSTITUTE

School of Pharmacy



Inclusion and exclusion criteria

Criteria	
Database	PubMed
Time period	1/1/2016 – 3/31/2021
Inclusion criteria	- At least 1 HEOR-related MeSH term - Developed or applied an ML technique for any purpose - Used data derived from wearable devices that were used outside the clinic
Exclusion criteria	- Non-original studies (e.g., review, commentary, etc) - Methods paper - Non-human studies - Non-English studies - Studies without full-text available - Studies using image, video, or –omics data



How we identified HEOR studies

Categories	MESH terms
Health Care Economics	Economics; Health Planning; Policy; Technology Assessment, Biomedical
Health Care Quality, Access and Evaluation	Delivery of Health Care; Health Services Research, Public Health Systems Research; Quality of Health Care; Epidemiologic Factors; Outcome and Process Assessment, Health Care; Patient Satisfaction; Program Evaluation; Technology Assessment, Biomedical; Quality Indicators, Health Care; Standard of Care
Health Services Administration	Delivery of Health Care; Disease Management; Patient Satisfaction; Quality of Health Care; Outcome and Process Assessment, Health Care; Program Evaluation; Quality Improvement; Quality Indicators, Health Care; School Mental Health Services
Population Characteristics	Demography



Data abstraction

General information

- PICOTS

Types of ML applications

- Types of disease areas
- Application purposes (forecasting vs. monitoring)
- Model outcomes (e.g., clinical event, costs, etc.)

Data used to train the ML models

- Types of wearable devices

ML methods and performance metrics

- Types of ML techniques used
- The form of the outcome variable (continuous or categorical)
- Performance metrics



Results

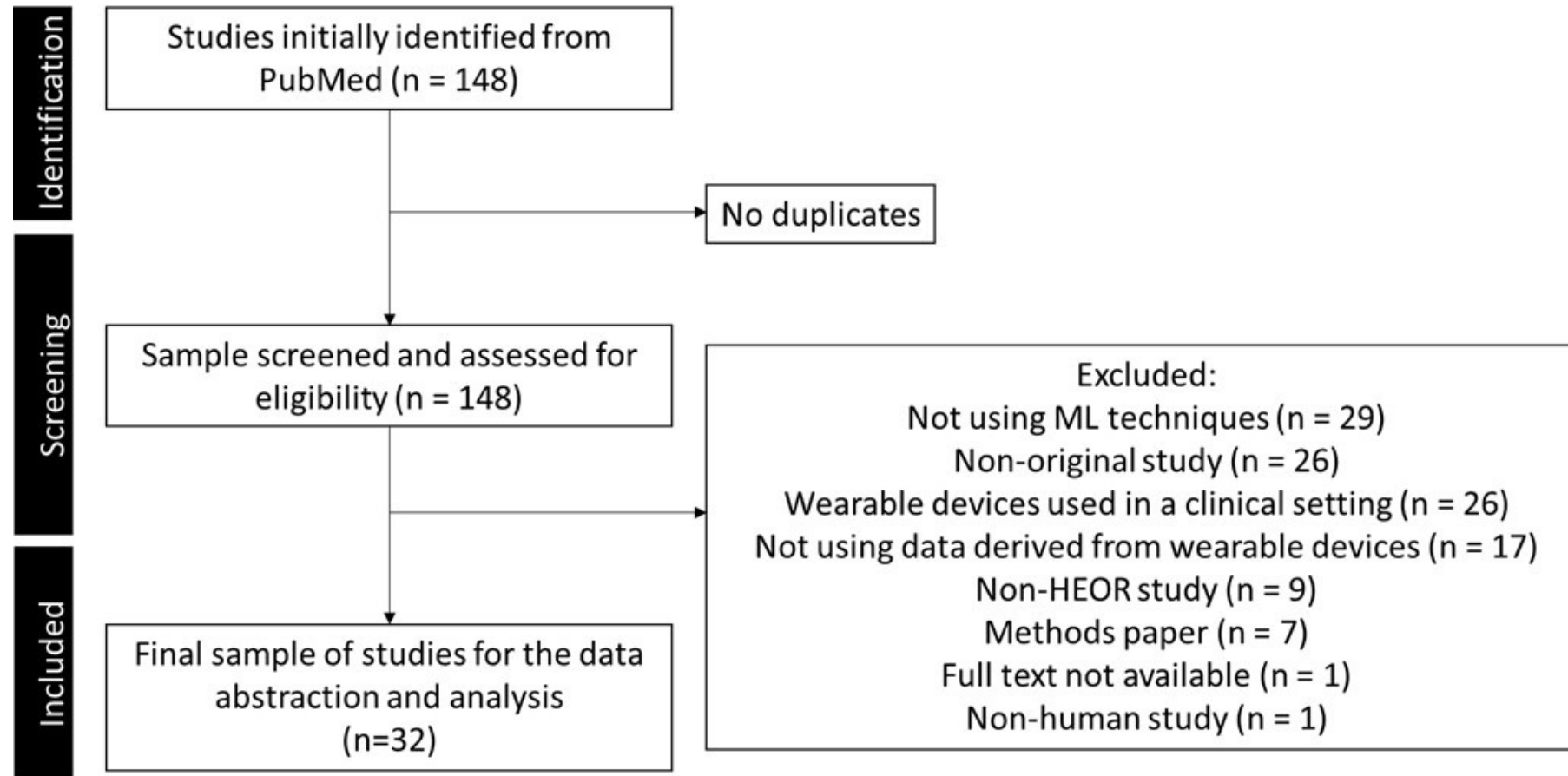


THE CHOICE INSTITUTE

School of Pharmacy



PRISMA (Preferred Reporting Items for Systematic Reviews) diagram



Types of disease areas

- > Not specific to any disease areas (17/32, 53%)
- > Parkinson's disease (3/32, 9%)
- > Cardiovascular diseases (3/32, 9%)
- > Others included mental disorders, cancer, respiratory diseases, etc.



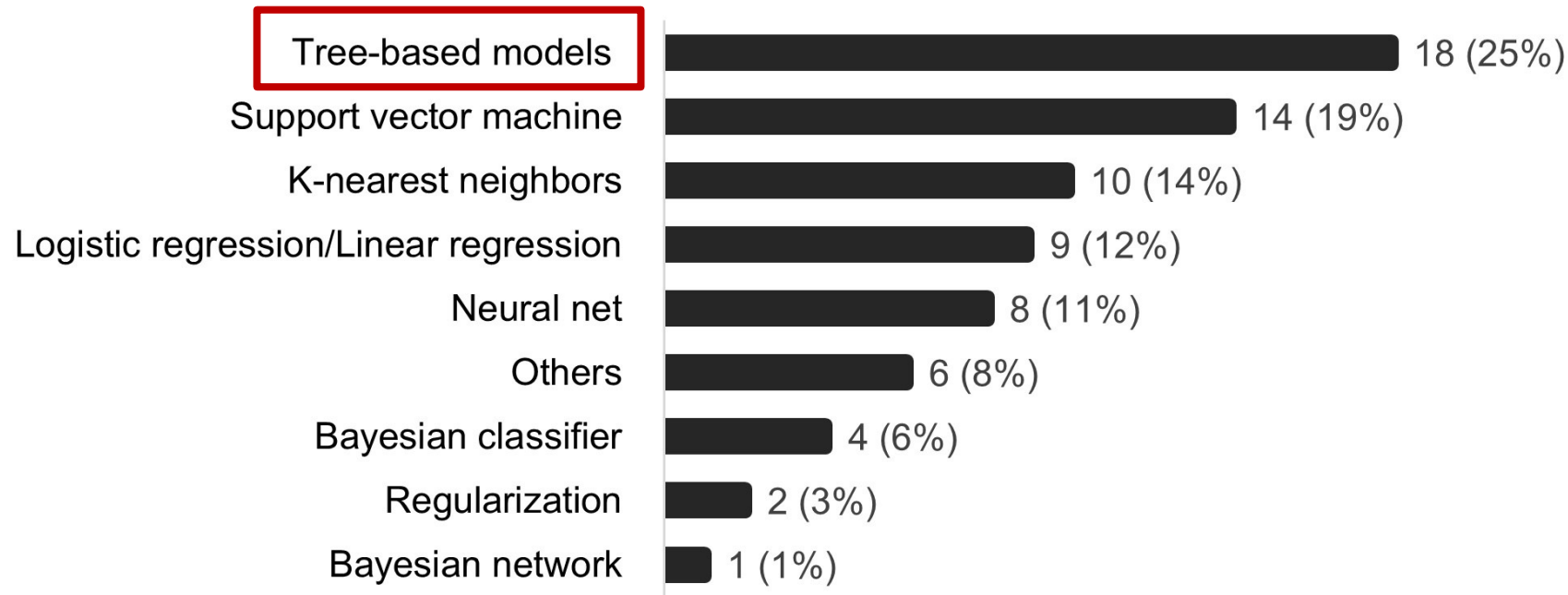
A wide range of outcomes examined, but mostly focused on monitoring

	Goal of application			
Model outcomes	Monitoring	Forecasting	All	Examples for model outcomes
General physical or mental health	6 (18%)	2 (6%)	8 (24%)	Stress or mental fatigue, physical functioning
Physical activities	7 (22%)	0 (0%)	7 (22%)	Activities of daily living, walking patterns
Clinical event or disease incidence	5 (16%)	1 (3%)	6 (19%)	Fall, seizure, freezing of gait
Disease progression or symptoms	4 (13%)	0 (0%)	4 (13%)	Amplitude and constancy of tremor, insomnia symptoms, mental disorder symptoms
Treatment patterns	3 (9%)	0 (0%)	3 (9%)	Medication intake, counts of pill medication
Treatment outcomes	0 (0%)	3 (9%)	3 (9%)	Post-operative complications, patient-reported outcomes following joint replacement
Healthcare resource utilization	0 (0%)	1 (3%)	1 (3%)	Hospital admission due to heart failure
Total	25 (78%)	7 (22%)	32 (100%)	

Non-ignorable number of studies that used data from devices without a medical purpose

Types of wearable devices	Number	Examples of model outcomes
Devices designed for medical purposes		
Medical devices with inertial sensors (e.g., Accelerometer or gyroscope)	10 (31%)	Activities of daily living, fall
Medical devices with ECG sensors (e.g., Chest strap sensor or band)	3 (9%)	Hospital readmission due to HF, stress detection
Telemeter	1 (3%)	Epileptic seizure
Others (e.g., multiple types or wearable devices combined, electrodermal activity device, pressure-sensing insoles)	5 (16%)	Balance function, pain, freezing of gait
Devices designed for non-medical purposes		
Smartwatches (e.g., Fitbit or Apple watch)	11 (34%)	Resting tremor, medication intake, emotional status
Smartphones (e.g., Phone app or audio function)	2 (6%)	Sensitivity to medication, fall
Total	32 (100%)	

Tree-based models were the most common type of ML technique



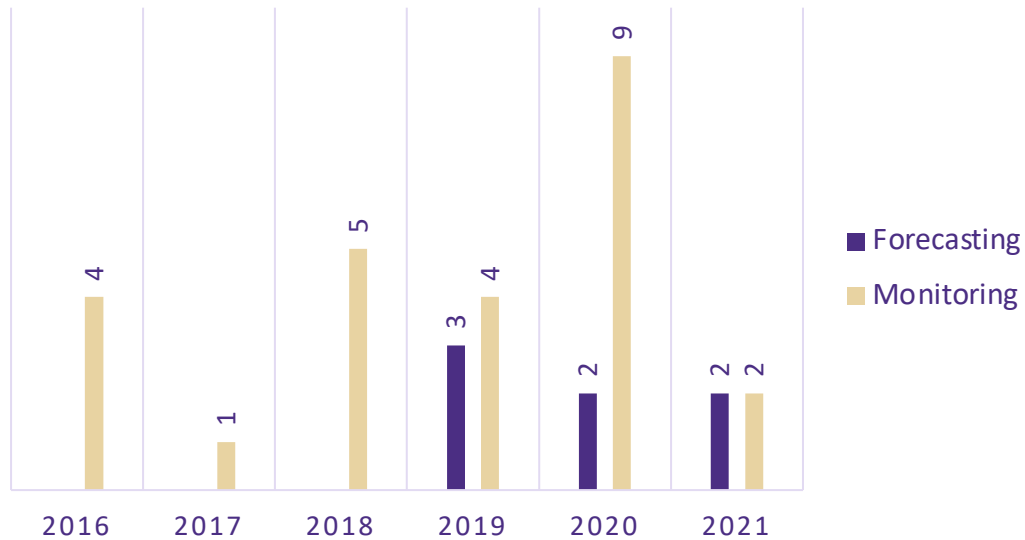
Accuracy was the most commonly used performance metric, followed by F1 score and AUROC

Performance metrics	Definition	
		All
Accuracy	$(TP + TN)/(TP + FP + FN + TN)$	14 (21%)
F1 score	$2 * (Recall * Precision)/(Recall + Precision)$	11 (17%)
AUROC	Area under the ROC curve	11 (17%)
Sensitivity or recall	$TP/(TP + FN)$	10 (15%)
Specificity	$TN/(TN + FP)$	8 (12%)
Precision or PPV	$TP/(TP + FP)$	5 (8%)
Sum of squares	$\sum_{i=0}^n (Y_i - \bar{Y})^2$	1 (2%)
Others	The area under the precision-recall curve (AUPRC), Average silhouettes, Kappa, difference with another variable of interest, Not mentioned	5 (7%)
Total		66 (100%)

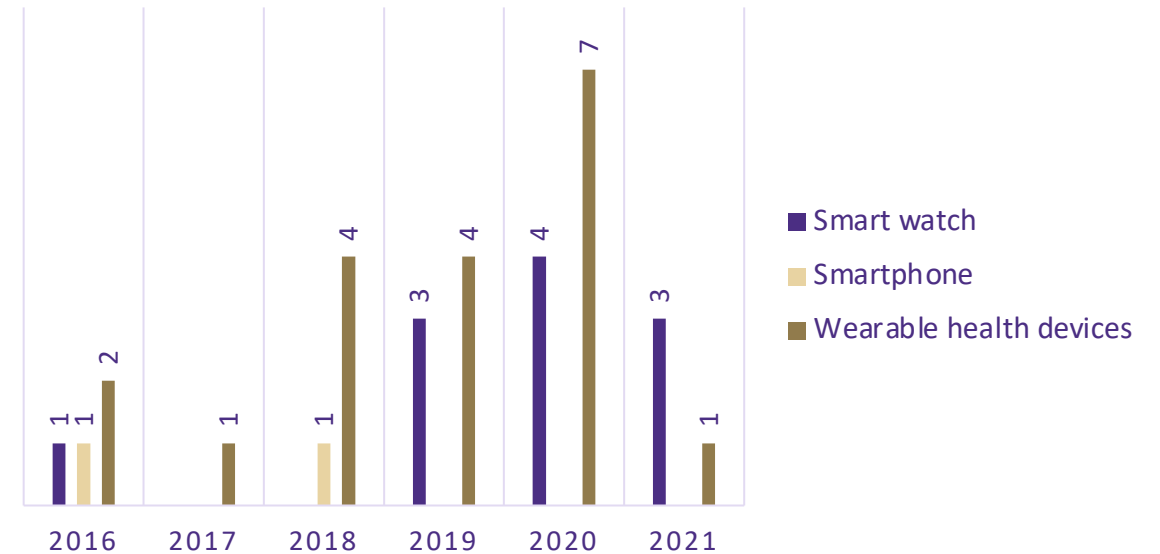


Time trend

Goal of application



Types of wearable data



Based on the descriptive analysis, studies that apply ML with forecasting purposes and those using smartwatches have become more common in recent years.



There's a potential for wearable devices to provide rich data for ML studies in HEOR

- > Although non-disease- or non-treatment-specific outcomes were common, there have been non-negligible number of studies specific to certain diseases or treatments
- > Studies that used devices without specific medical purposes accounted for approximately a half, and its number has been increased recently
- > Challenges: data quality, selection bias, security and privacy

